

# **Window-Based Applications of TRC Databases: Structure and Internet Distribution**

X. Yan<sup>a\*</sup>, Q. Dong<sup>a</sup>, M. Frenkel<sup>a</sup>, K. R. Hall<sup>a, b</sup>

<sup>a</sup>Thermodynamics Research Center, Texas A&M University System,  
College Station, TX 77843-3111, USA

<sup>b</sup>Chemical Engineering Department, Texas A&M University,  
College Station, TX 77843-3122, USA

\*To whom correspondence should be addressed

## ABSTRACT

Thermophysical and thermochemical property data of organic compounds provides vital background information for scientific and engineering communities in both academic and industrial environment. The Thermodynamic Research Center (TRC) has collected an enormous amount of property data over the past 55 years. New advances in computer technology enable relatively large databases to have friendly ‘communications’ based upon a PC Windows platform. Here, we discuss the structure, upgrade procedures, and the mechanisms for distribution over local PC networks and the Internet with emphasis on 3 databases: TRC COMPOUNDS (Wincmp), SOURCE (WinSource), the archive of raw experimental data, and TRC TABLE (WinTable) the TRC collection of recommended data.

Key words: thermophysical property, database, Windows, experimental data, recommended data, Internet, compilation, evaluation.

## 1. INTRODUCTION

The thermodynamics Research Center (TRC) has performed physico-chemical property collection, critical evaluation and data management for more than 55 years. For the past 20 years, TRC has developed a number of electronic databases for thermophysical properties [1]. Among these databases, SOURCE and TABLE are the two major ones. SOURCE [2] has existed for 14 years and is an archive of original, experimentally measured values of thermophysical and thermochemical data covering over 100 properties of pure compounds, binary and ternary systems, and chemical reactions (Fig. 1). Recently, TRC has rebuilt the SOURCE under the ORACLE client/server database system (ORACLE 8 Enterprise Edition) assuring robust maintenance and consistent data entry procedures [3].

The COMPOUNDS database is a subset of SOURCE representing a compound identification block. It contains Chemical Abstract Service Registry Numbers (CASRN), Chemical Formulas and molar masses for more than 113,000 compounds with more than 218,000 alternative chemical names (Fig. 1).

The TABLE database is essentially an electronic version of TRC Tables-Hydrocarbons and TRC Tables-Non-Hydrocarbons [4, 5] published by TRC since 1942, and updated 4 times yearly. TABLE contains 31 recommended thermophysical properties for 7,446 compounds also in the ORACLE format residing on a DEC  $\alpha$  Workstation.

To provide users with easy access and multiple options to extract, display and distribute data, TRC has developed new Window-based versions of TRC databases. TRC maintains three databases with the ACCESS database support system having retrieval functions

written in Visual Basic. Efficient procedures exist to update these databases by converting data from ORACLE to ACCESS format, providing easy access through Windows GUI interfaces to display numerical data in tabular and graphic forms, to fit data with selected equations, and to distribute data through the Internet.

## 2. DESIGN AND PRODUCTION

Principal design and production schema for all 3 databases (Wincmp, WinSource, and WinTable) are similar as illustrated by Fig. 2. All 3 databases are essentially read-only versions of the TRC ORACLE in-house databases, SOURCE or TABLE. Converting the data from the ORACLE database into ACCESS database through Open DataBase Connectivity (ODBC) – a common gateway for all kinds of databases - populates these databases. In turn, data entry into TRC ORACLE databases is possible either through the interactive Data Entry Form supported by ORACLE Developer 2000 or using Batch Data Input Mode. The data stored in the MS ACCESS format can be retrieved, correlated, fitted to the selected equations and represented in both tabular and graphical forms using the interfaces designed using Visual Basic. The tables of MS ACCESS database can also be converted into an ASCII file for further use in other applications such as process simulation. Both ORACLE and MS ACCESS databases can be distributed *via* the Internet through the ORACLE or MS Web Server. Two files built with ChemDraw characterize each compound. One file is an image corresponding to the particular structural formula while the other file contains the Connectivity Table.

### 3. Wincmp - TRC COMPOUND DATABASE FOR PC WINDOWS

We have produced and upgraded the ACCESS database for Wincmp from the compound identification block of the ORACLE SOURCE. Other TRC Windows databases use the retrieval functions and formats developed to represent output from Wincmp. Using Wincmp, it is possible to retrieve compound identification information by CASRN, chemical formula, formula fragment, chemical name, chemical name fragment, and any combination of these search criteria as well as in batch mode. The formula fragment and name fragment functions permit a narrower search in case the net formula and full chemical name are unknown or doubtful. These options are also valuable if the search is related to the group of compounds characterized by the same formula fragment or name fragment. The search results obtained using the functions formula fragment or name fragment can be saved into a data subset for further search.

### 4. WinSource - TRC SOURCE DATABASE FOR PC WINDOWS

We have produced and upgraded the ACCESS database for WinSource from the ORACLE version of SOURCE. Data entry for the ORACLE SOURCE is possible using both the Data Entry Form option (mostly in-house) and the Batch Entry Mode (to incorporate data files supplied by external contributors, Fig. 2). The available retrieval functions from WinSource fall into five groups: compound identification, pure compound properties, binary and ternary systems properties, chemical reaction properties, and references.

The compound identification block in WinSource is identical to Wincmp except it contains only those compounds which have numerical data (over 15,800 compounds, Fig. 1).

Pure compound property data records cover constants, property at fixed conditions, and 1- and 2-variable properties, totaling 50 properties: adiabatic compressibility, [1/kPa], isothermal compressibility, [1/kPa], compressibility factor, [PV/RT], critical pressure, [kPa], critical temperature, [K], critical molar volume, [m<sup>3</sup> mol<sup>-1</sup>], critical density, [kg m<sup>-3</sup>], molar density, [mol m<sup>-3</sup>], specific density, [kg m<sup>-3</sup>], molar volume, [m<sup>3</sup> mol], specific volume, [m<sup>3</sup> kg], self-diffusion coefficient, [m<sup>2</sup> s<sup>-1</sup>], thermal diffusivity, [m<sup>2</sup> s<sup>-1</sup>], heat capacity ratio,  $C_p/C_v$ , heat capacity at constant pressure, [J K<sup>-1</sup> mol<sup>-1</sup>], heat capacity at constant pressure per unit volume, [J m<sup>-3</sup> K<sup>-1</sup>], heat capacity at saturation, [J mol<sup>-1</sup> K<sup>-1</sup>], heat capacity at constant volume, [J mol<sup>-1</sup> K<sup>-1</sup>], enthalpy of transition, [kJ mol<sup>-1</sup>], enthalpy of vaporization at saturation, [kJ mol<sup>-1</sup>], fugacity coefficient,  $f/P$ , pressure at phase equilibrium, [kPa], vapor pressure, [kPa], refractive index for isotropic phase, Na D-line, refractive index of isotropic phase, other wave length, speed of sound, [m s<sup>-1</sup>], surface tension, [N m<sup>-1</sup>], triple point, [K], boiling point, [K], normal boiling point, [K], normal freezing point, [K], Gibbs energy at 1 bar, [kJ mol<sup>-1</sup>], reduced Gibbs energy,  $(G-H_0)/T$ , [J K<sup>-1</sup> mol<sup>-1</sup>], enthalpy, [kJ mol<sup>-1</sup>], reduced enthalpy,  $(H-H_0)/T$ , [J K<sup>-1</sup> mol<sup>-1</sup>], pressure coefficient of enthalpy, [kJ mol<sup>-1</sup> kPa<sup>-1</sup>], thermal conductivity, [W m<sup>-1</sup> K<sup>-1</sup>], thermal pressure coefficient, [kPa K<sup>-1</sup>], entropy, [J K<sup>-1</sup> mol<sup>-1</sup>], Joule-Thompson coefficient, [K kPa<sup>-1</sup>], internal energy, [kJ mol<sup>-1</sup>], coefficient of expansion, [K<sup>-1</sup>], cryoscopic constant, [K], second acoustic virial coefficient, [m<sup>3</sup> mol<sup>-1</sup>], second virial coefficient, [m<sup>3</sup> mol<sup>-1</sup>], third

virial coefficient, [ $\text{m}^6 \text{mol}^{-2}$ ], third acoustic virial coefficient, , [ $\text{m}^6 \text{mol}^{-2}$ ], viscosity, [ $\text{Pa s}$ ], kinematic viscosity, [ $\text{m}^2 \text{s}^{-1}$ ].

The retrieval functions developed for WinSource provide a variety of options for users to search all or any combination of properties data values available for a compound as well as particular property data for all or a selected group of compounds.

After having selecting a particular compound and property for the search, the results appear in both textual and tabular forms and can be saved into a user-designated file. If the retrieved set contains more than 2 data points, the results of the search are available in graphical mode with options available for representation of uncertainties, selection of a particular range of the variables and properties and particular user-designated data points. For constant and fixed- condition data, WinSource provides display of the numerical data, deviation plots (as an option), average value and average deviation. An example of the display from WinSource for a critical temperature search for methanol is the Fig. 3.

WinSource provides a collection of more than 330,000 data points covering around 100 thermophysical properties for more than 8,500 binary mixtures. It also includes all thermodynamic and transport property data collected by TRC during the development of the Design Institute for Physical Property Data (DIPPR) Project 882 [6]. To select a mixture for data retrieval, the user first selects one component of the binary mixture. WinSource then displays a list of all second components of the binaries, which, in combination with first component, represent mixtures characterized by numerical data in the database. Upon selection of the second binary component from this list, WinSource displays the list of the available properties for the chosen mixture in the particular

thermodynamic phase. Having defined the combination mixture/property, the search proceeds using retrieval functions similar to those developed for pure compound data retrievals. Fig. 4 is an example display of the results for a search of surface tension data for the binary mixture methanol-carbon tetrachloride.

WinSource contains about 10,000 data points covering 20 properties for 555 ternary mixtures. The ternaries can be selected directly from the complete list of the ternary mixtures available or by stepwise procedure similar to that for binaries. In addition to the options used to display the results of the data search for binary mixtures, ternary mixture data can be displayed in the triangle-type format when the variables (property) are mole (mass) fractions of the components.

WinSource contains thermodynamic data for more than 2500 chemical reactions. These data reside in two different tables characterizing those obtained for reactions with initial thermodynamic state change (usually irreversible chemical reactions studied by bomb and reaction calorimetry) and those characterizing chemical reactions at equilibrium. Most of the first type data encompass such properties as change of internal energy, and heat and enthalpy of reaction. Various types of equilibrium constant properties (usually as a function of temperature at constant pressure) characterize the second type data. The collection of equilibrium data in WinSource essentially covers all the experimental data of this type available including those previously selected by Frenkel *et al* [7]. Twelve properties are available from the reaction block in WinSource: heat capacity at constant pressure,  $[J K^{-1} mol^{-1}]$ , enthalpy,  $[kJ mol^{-1}]$ , internal energy,  $[kJ mol^{-1}]$ , heat of combustion at constant volume,  $[J g^{-1}]$ , Gibbs energy,  $[kJ mol^{-1}]$ , entropy  $[J K^{-1} mol^{-1}]$ , volume,  $[m^3$



mol<sup>-1</sup>], apparent equilibrium constant in terms of molarities, apparent equilibrium constant in terms of molalities, apparent equilibrium constant in terms of partial pressures in the gas phase, thermodynamic equilibrium constant for the specified standard state, and apparent equilibrium constant in terms of mole fraction.

All the reactions are subdivided into 18 types: addition of unsaturated compounds to other compounds, combustion with oxygen, combustion with other compounds or elements (Cl<sub>2</sub>, F<sub>2</sub>, *etc.*), homonuclear dimerization, esterification, exchange of alkyl groups, exchange of hydrogen with other groups, formation from elements, halogenation (addition or replacement), addition of water to liquid or solid compounds to produce a hydrate, other reactions with water, hydrogenation (addition of hydrogen to unsaturated compounds), hydrohalogenation, structural isomerization, stereo isomerization, neutralization (acid plus base), oxidation with oxygen (not complete combustion), other reactions (not in any other class).

WinSource provides thermodynamic data for chemical reactions by performing searches based upon reaction type, property and one component in the reaction mixture. The reaction block in WinSource contains more than 3,000 components. Structure files created using ChemDraw (Fig. 2) characterize all components and permit display of chemical reactions in a structural form. The equilibrium data as a function of temperature can be represented graphically using normal or logarithmic coordinates. Fig. 5 presents an example display of the results for an equilibrium data search for the hydrogenation of naphthalene.

WinSource provides additional retrieval functions for information related to the original sources of experimental data totaling of more than 81,000 references. References can be searched by 1) reference code consisting of the year of the publication, 3 or 6 characters corresponding to the first 3 letters of the author's last name or of the first and the second authors' last names (when the reference is authored by more than 1 author) and additional sequence number in case if the combination year/author(s) is not unique; 2) author name or name fragment; 3) given period of time with the option to use author name or name fragment as additional search criteria; 4) given period of time and type of the property data reported.

#### 5. WinTable - TRC TABLE DATABASE FOR WINDOWS

The ACCESS database for WinTable comes from the ORACLE version of TABLE database. This database stores recommended data for a selected list of thermodynamic properties of pure organic compounds. It is populated with data using specially-developed procedures which take data directly from input files used to update the on-going loose-leaf serial publications: TRC Thermodynamic Tables-Hydrocarbons [4] and TRC Thermodynamic Tables-Non-Hydrocarbons [5]. The properties currently covered in WinTable are: properties in the ideal gas state (heat capacity, entropy, reduced enthalpy, reduced Gibbs energy, enthalpy and Gibbs energy of formation), thermodynamic properties in the condensed state (heat capacity, entropy, reduced enthalpy, reduced Gibbs energy), critical properties (temperature, pressure, volume, density, compression factor), adiabatic compressibility, enthalpy of combustion and formation in various phases, enthalpy and temperature of phase transition (crystal-crystal, crystal-liquid, liquid-gas,

crystal-gas), transport properties (viscosity, thermal conductivity, speed of sound), surface tension, refractive index, specific density, second virial coefficient, normal boiling and freezing point, triple point, and vapor pressure.

The retrieval functions for WinTable are similar to those available for the pure compound block in WinSource. In addition, WinTable has particular tables designed to store the coefficients of fitting equations used for interpolation and extrapolation (when possible).

## 6. FUTURE DEVELOPMENT

The availability of critically evaluated and recommended thermodynamic and thermophysical property data is crucial for great variety scientific and engineering applications. Therefore, in the near future TRC intends to expand significantly the amount of recommended data available in electronic form beyond the coverage of TRC Thermodynamic Tables (Fig. 6). The combination of recommended data currently available from WinTable with data previously evaluated by TRC for a variety of properties (density [8-14], vapor pressure [15, 16], ideal gas properties [17]) provide data in the form of equation coefficients along with data generated directly from SOURCE using the principle of dynamic compilation [18]. Predicted values generated using group contribution methods based upon the connectivity table files created with chemical structure image files (Fig. 2) can fill existing ‘gaps’ in available experimental data. Currently, the structure information for more than 4,000 compounds exists in a database. Molecular adjacency matrix, distance matrix, topological index, group types for group contribution methods and some molecular descriptors result from this information. This

comprehensive collection of the recommended data with thermodynamic consistency create the foundation for a new database, SELECT, whose direct output applies to numerous applications.

## 7. DISTRIBUTION VIA THE INTERNET

The benefits of developing and deploying applications over the World Wide Web become apparent when combining databases and Web technology. Internet applications that count on fast and reliable database systems are deployed and maintained on centralized application servers, from which they are downloaded to the end-user's Web browser at runtime. This distribution method reduces the time, cost, and complexity of deploying applications to a large or geographically dispersed end-user base, all without installing application software the individual desktop machines.

WEBTRC is an ongoing project at TRC concerned with building a dynamic web site for thermophysical property data distribution. The site uses a three-tier architecture. The first tier is the Web browser. The middle tier consists of Oracle Developer Server and Microsoft Internet Information Server running on a Microsoft Windows NT server. The third tier is the database tier that includes Oracle 8I running on a DEC  $\alpha$  workstation. The goal of the project is to enable end-users to retrieve data interactively or order data products through the browser and Internet access. Another key role of this site is to provide data centers and individual experts around the world with the ability to input experimental data directly into a mirror database of SOURCE, Source Feeds, through a standard data entry form.

Completion of the first phase in this project includes:

- (1) A distributed data collection system consisting of Source Feeds and the Data Entry Form. Incoming data from the Source Feeds through the Data Entry Form are analyzed by the database rules established at TRC as SQL triggers that generate an error report on the incoming data. TRC staff tests and checks for erroneous data in the mirror database before loading data into SOURCE.
- (2) A web retrieval system consisting of the Compound database, the Table database, and the Source database. The anticipated goal is to provide end-users with identical feature-rich functions such as the standard desktop versions described in this paper.

Web-based architecture represents a shift back to a server-centric deployment model. This model enables an effective and efficient distribution mechanism without relying upon individual client configurations. A Web database application process flow designed for the WEBTRC site appears in Fig. 7. To start and run any databases from WEBTRC, end-users can deploy their selected Web browser to access an URL. The URL corresponds to a web page on the Web server. An HTML page and Form client scripts (for an Oracle process) or other client-site scripts such as Active Server Page scripts (for a Microsoft process) are generated accordingly and then downloaded to the user's browser. These functions contact the Form runtime engine or ASP server scripts to initiate a new database process or to join an existing process. The Web server engine is then in charge of communicating with databases on the database server. The engine sends the retrieved results back to the client's browser. This Internet distribution schema allows connection of a large number of users

worldwide to the database collection available at TRC for customized or standard applications.

## REFERENCES

1. TRC Databases for Chemistry and Engineering – Catalogue : (Thermodynamics Research Center, College Station TX, 2000), see also website: <http://trcweb.tamu.edu>
2. Documentation for the TRC Source Database (Thermodynamics Research Center, College Station, TX, 1999), 207 pp. An abridged version can be downloaded from the TRC Web site, [trcweb.tamu.edu](http://trcweb.tamu.edu).
3. M. Frenkel, Q. Dong, R. C. Wilhoit, K. R. Hall, TRC SOURCE database: a unique tool for automatic production of data, presented at the 14-th Symposium on Thermophysical Properties, Boulder, CO, June 25-30 (2000)
4. TRC Thermodynamic Tables-Hydrocarbons (TRC, College Station, TX, 1952-2000)
5. TRC Thermodynamic Tables-Non-Hydrocarbons (TRC, College Station, TX, 1956-2000)
6. K. N. Marsh, Q. Dong, A. K. R. Dewan, Transport properties and related thermodynamic data of binary mixtures, Part 1-5 (AIChE, New York, 1993-1997)
7. M. L. Frenkel, G. Ya. Kabo, G. N. Roganov, Thermodynamic properties of isomerization reactions (Hemisphere, Washington, DC, 1993)
8. Z.-Y. Zhang, M. Frenkel, K. N. Marsh, R. C. Wilhoit, Enthalpies of Fusion and Transition of Organic Compounds: Landolt-Bornstein, New Series, Vol. IV/8A, (Springer-Verlag, Berlin. 1995)
9. R. C. Wilhoit, K. N. Marsh, X. Hong, N. Gadalla, M. Frenkel, Densities of Aliphatic Hydrocarbons: Alkanes, Landolt-Bornstein, New Series, Vol. IV/8B, (Springer-Verlag, Berlin. 1996)

10. R. C. Wilhoit, K. N. Marsh, X. Hong, N. Gadalla, M. Frenkel, Densities of Aliphatic Hydrocarbons: Alkenes, Alkadienes, Alkynes and Miscellaneous Compounds, Landolt-Bornstein, New Series, Vol. IV/8C, (Springer-Verlag, Berlin. 1996)
11. R. C. Wilhoit, X. Hong, M. Frenkel, K. R. Hall, Densities of Monocyclic Hydrocarbons, Landolt-Bornstein, New Series, Vol. IV/8D, (Springer-Verlag, Berlin. 1997)
12. R. C. Wilhoit, X. Hong, M. Frenkel, K. R. Hall, Densities Aromatic Hydrocarbons, Landolt-Bornstein, New Series, Vol. IV/8E, (Springer-Verlag, Berlin. 1998)
13. R. C. Wilhoit, X. Hong, M. Frenkel, K. R. Hall, Densities Polycyclic Hydrocarbons, Landolt-Bornstein, New Series, Vol. IV/8F, (Springer-Verlag, Berlin. 1999)
14. M. Frenkel, X. Hong, R. C. Wilhoit, K. R. Hall, Densities of Alcohols, Landolt-Bornstein, New Series, Vol. IV/8G, (Springer-Verlag, Berlin. 2000)
15. J. Dykyj, J. Svoboda, R. C. Wilhoit, Frenkel, K. R. Hall, Vapor Pressure and Antoine Coefficients for Hydrocarbons, Sulfur-, Selenium-, Tellurium- and Halogen-Containing Organic Compounds, Landolt-Bornstein, New Series, Vol. IV/20A, (Springer-Verlag, Berlin, 1999)
16. J. Dykyj, J. Svoboda, R. C. Wilhoit, Frenkel, K. R. Hall, Vapor Pressure and Antoine Coefficients for Oxygen - Containing Organic Compounds, Landolt-Bornstein, New Series, Vol. IV/20B, (Springer-Verlag, Berlin, 2000)
17. M. Frenkel, G. J. Kabo, K. N. Marsh, G. N. Roganov, R. C. Wilhoit, Thermodynamics of Organic Compounds in the Gas State, Vol. 1, 2 (TRC: College Station, TX, 1994)
18. R. C. Wilhoit and K. N. Marsh, Int. J. Thermophys. 10:247 (1999)



## FIGURE CAPTIONS

Fig. 1. Statistical characteristics of TRC SOURCE database (see up-to-date information on TRC Web site <http://trcweb.tamu.edu>)

Fig. 2. Production lines for TRC Window-based databases

Fig. 3. Example display of the critical temperature data search for methanol from WinSource

Fig. 4. Example display of the surface tension data search for the binary mixture methanol-carbon tetrachloride from WinSource

Fig. 5. Example displaying results of the equilibrium constant data search for the hydrogenation of naphthalene from WinSource.

Fig. 6. Schema of the potential expansion of the WinTable.

Fig. 7. Web database process flow.

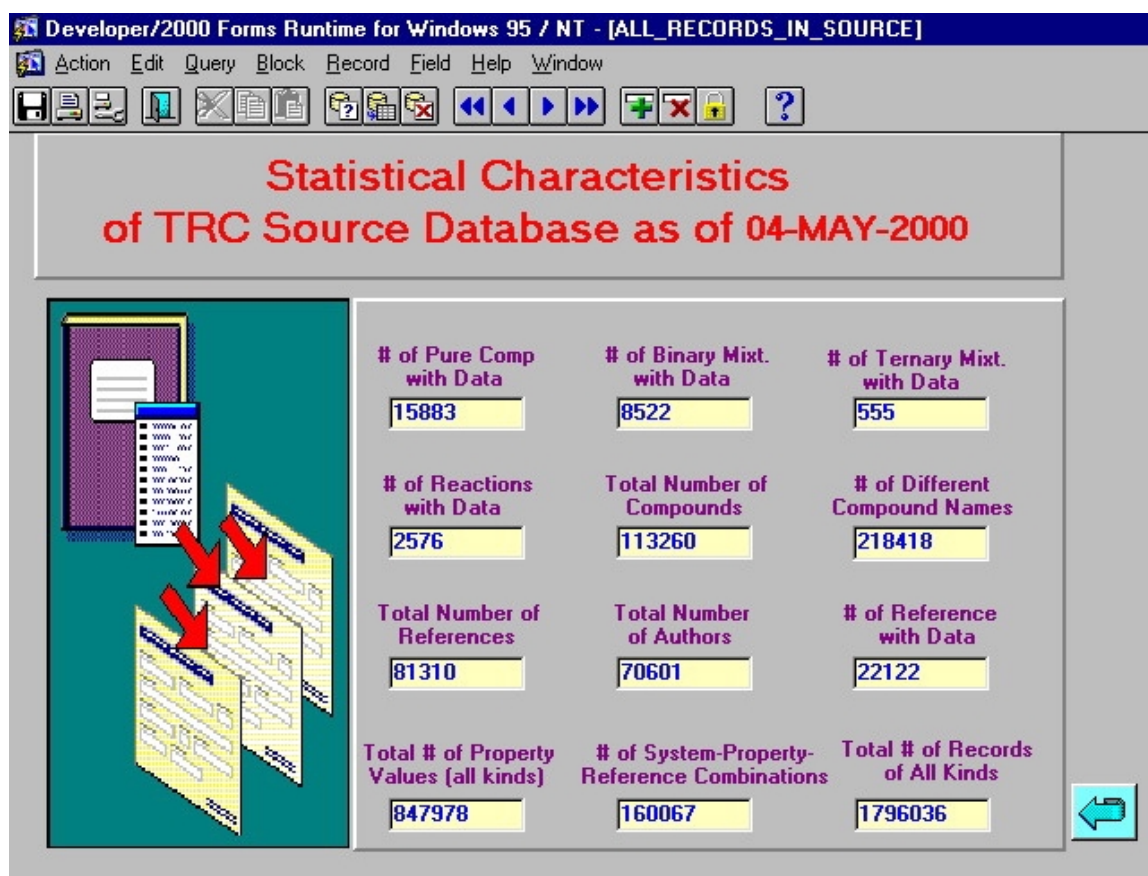


Fig. 1

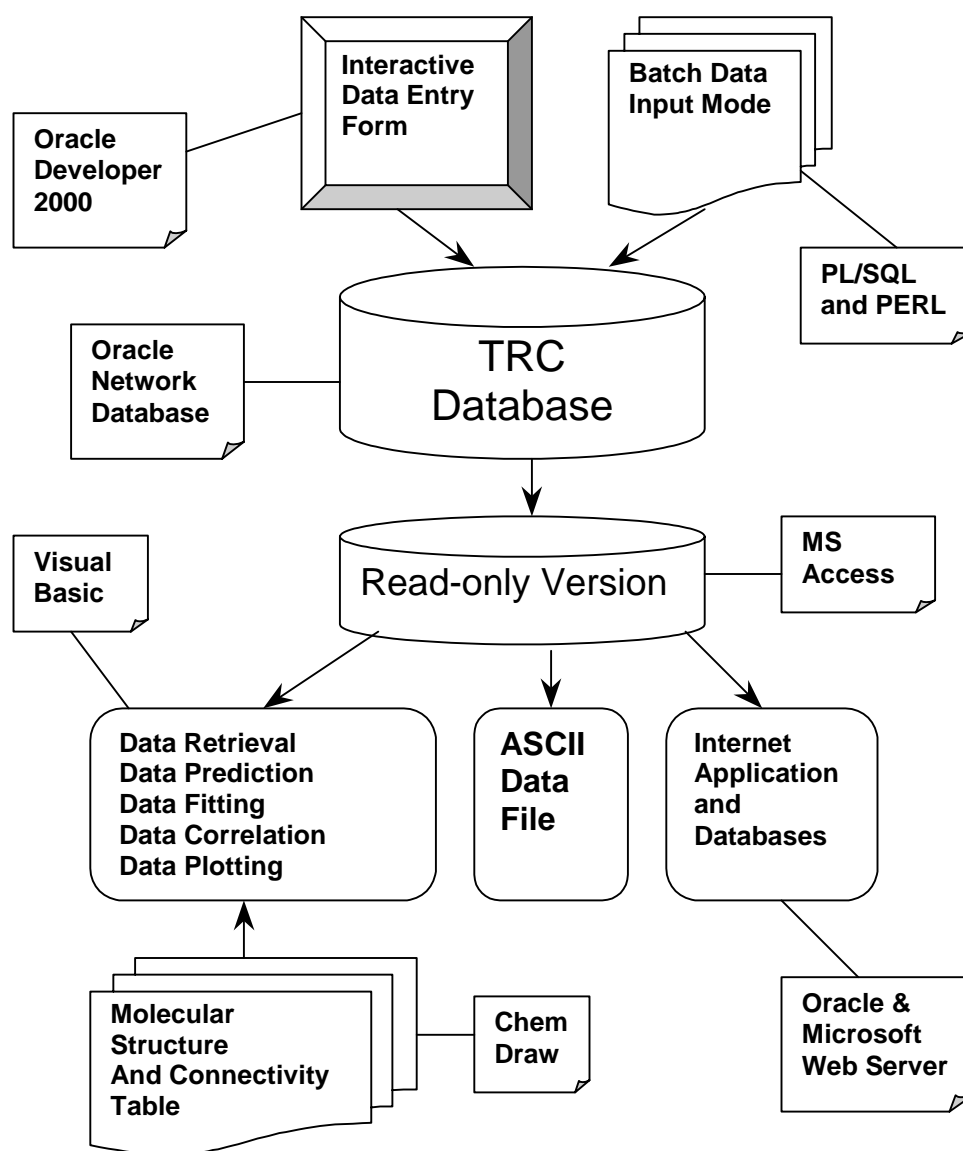


Fig. 2

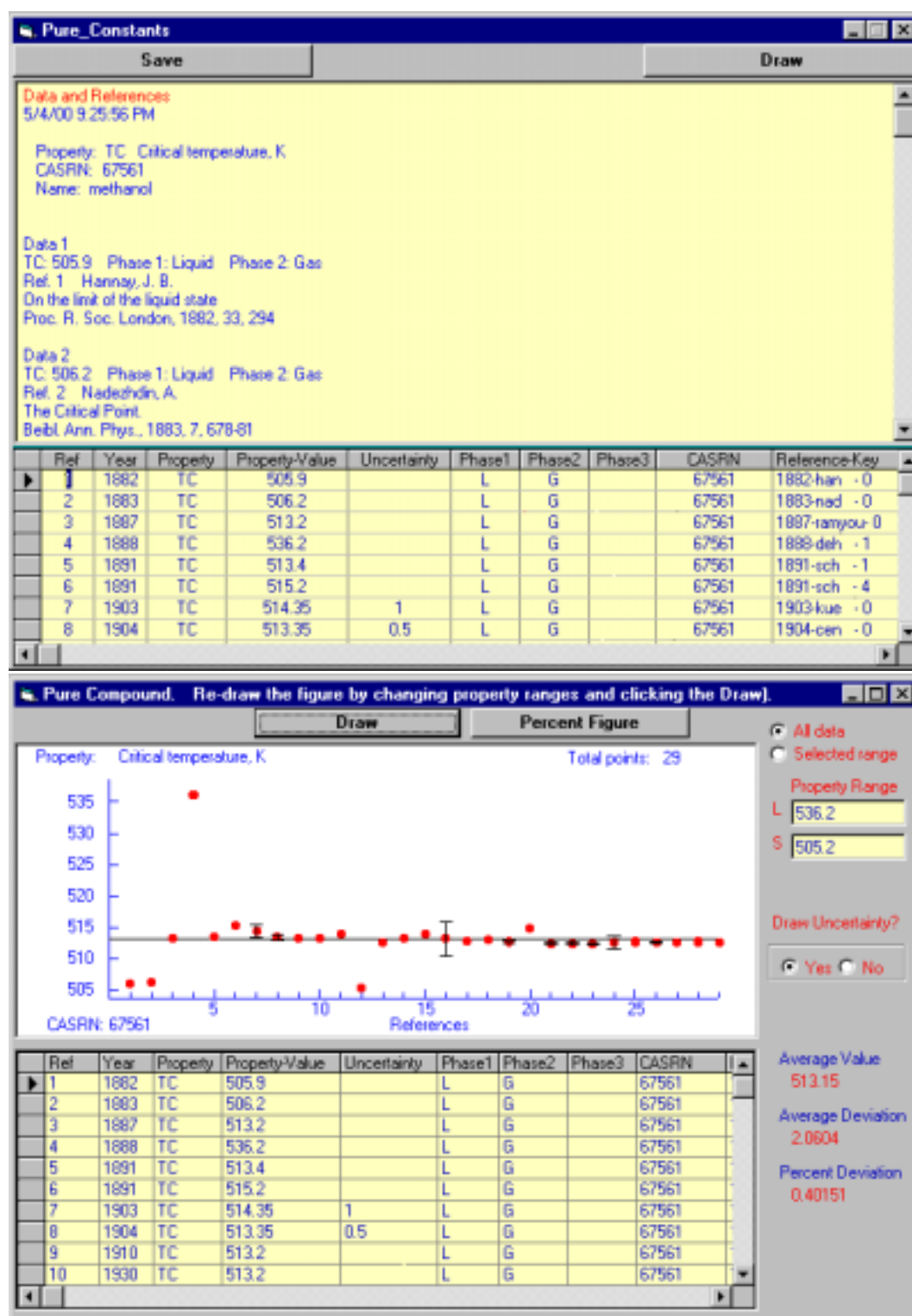


Fig. 3

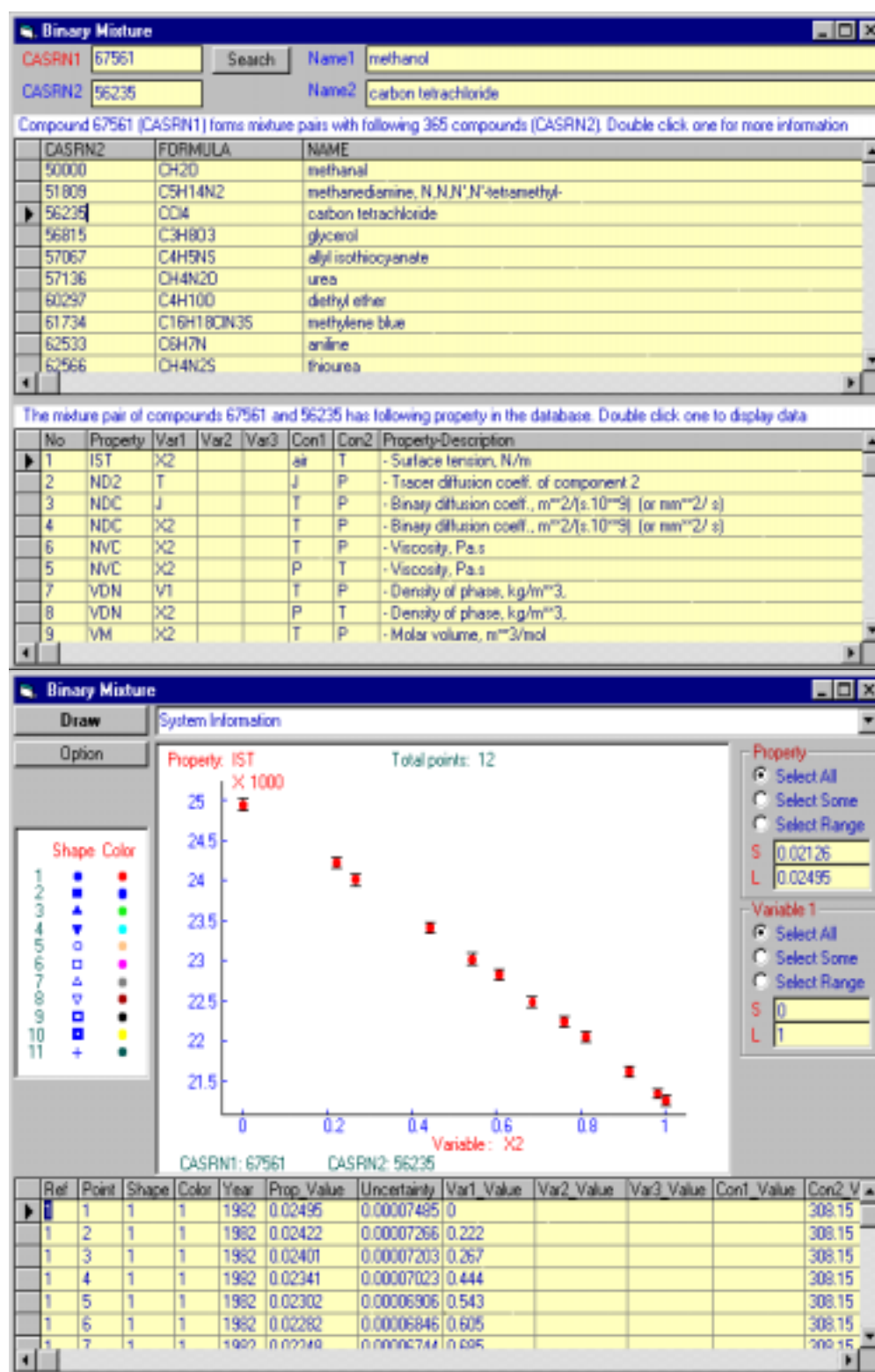


Fig. 4

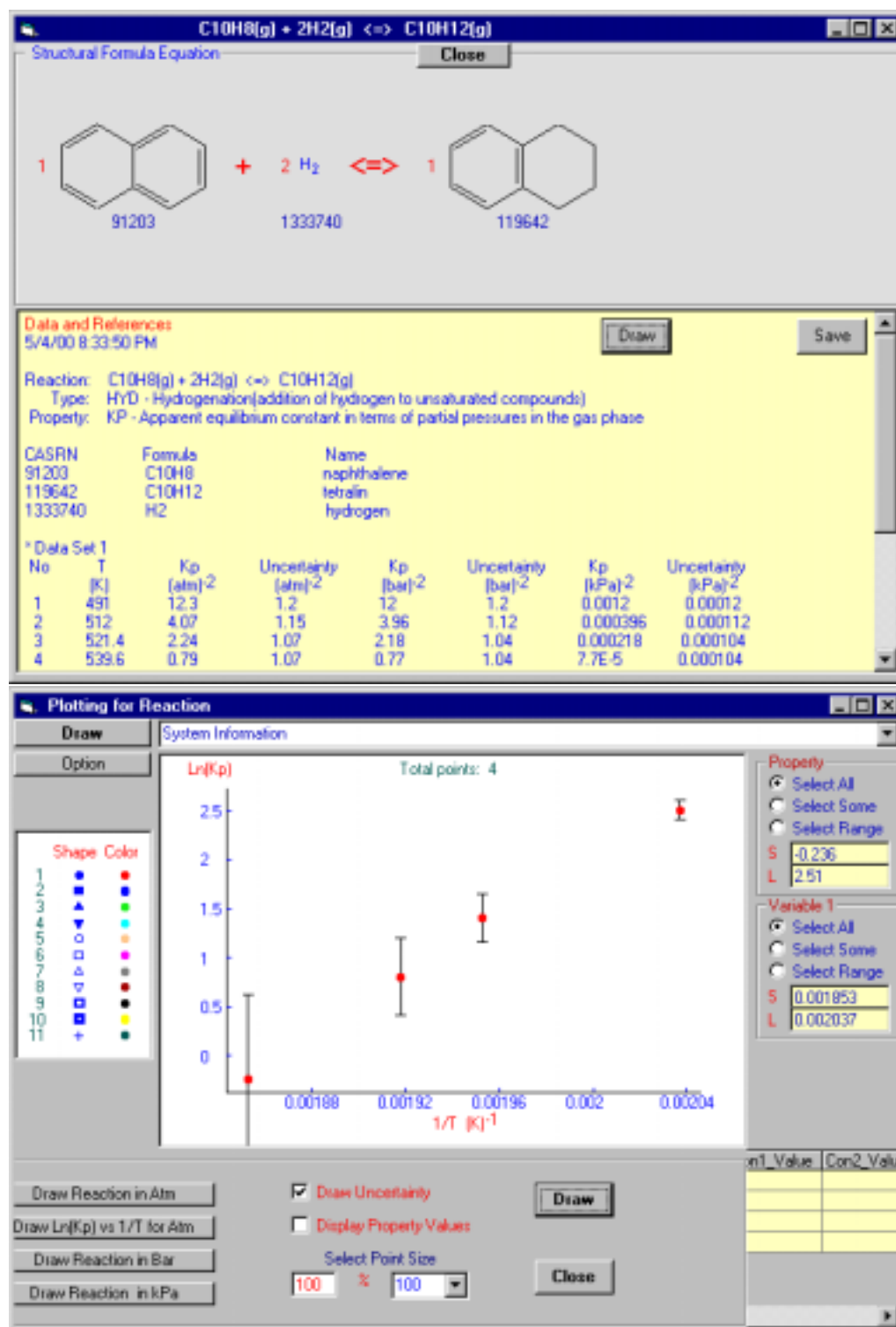


Fig. 5

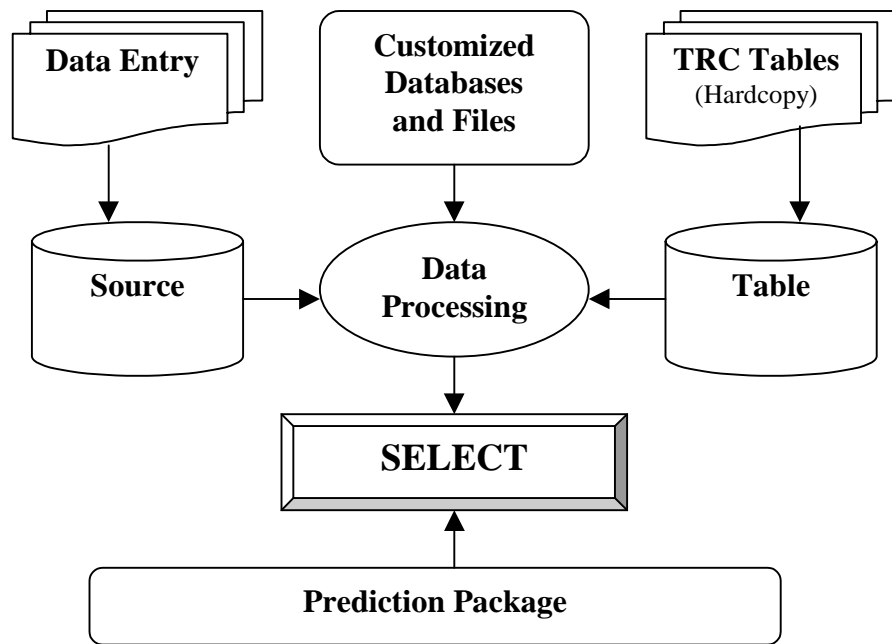


Fig. 6

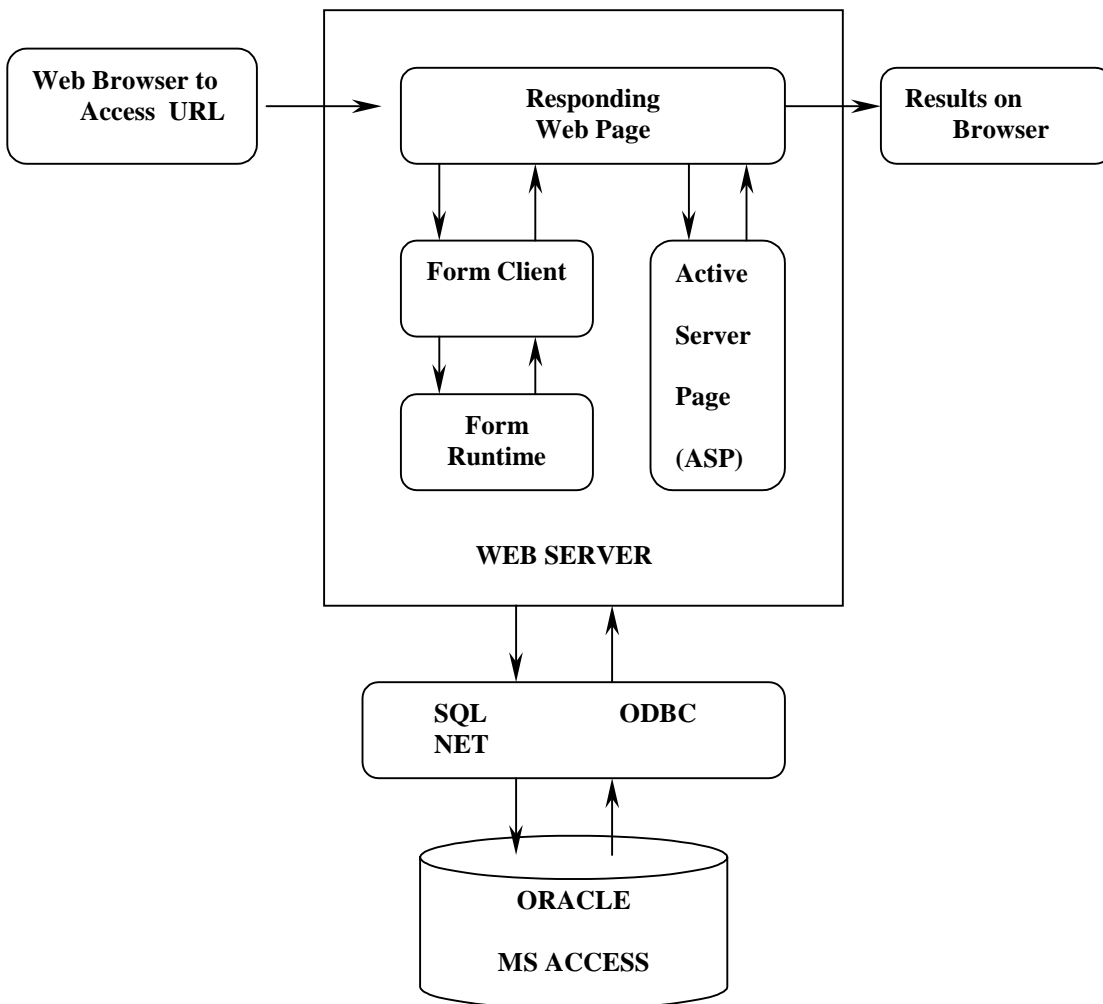


Fig. 7